

Tartu Ülikool
Loodus- ja tehnoloogiateaduskond
Ökoloogia ja Maateaduste Instituut
Geograafia osakond

Bakalaureusetöö geoinformaatikas

**Ürituste automaatne tuvastamine passiivse mobiilpositsioneerimise
andmetelt**

Taavi Pipar

Juhendaja: MSc Erki Saluveer

Kaitsmisele lubatud:

Juhendaja:

Osakonna juhataja:

Tartu 2014

Sisukord

Sissejuhatus	3
1. Teoreetilised lähtekohad	4
1.1 Ürituse definitsioon	4
1.2 Ürituste uuringute lähteandmed.....	5
1.3 Ürituste liigitamine	6
2. Andmed ja metoodika	11
2.1 Mobiilpositsioneerimise andmed.....	11
2.2 Andmetöötlus.....	11
2.3 Ürituste liigitamine	17
3. Tulemused	18
3.1 Automaatselt tuvastatud üritused.....	18
3.2 Aja ja asukoha tuvastamine	22
4. Arutelu.....	27
4.1 Ürituste tuvastamise täpsus	27
4.2 Ürituste tuvastamisel tekkivad vead	27
5. Kokkuvõte	30
Tänuavaldused.....	31
Summary	32
Kasutatud kirjandus.....	33

Sissejuhatus

Nõudlus ürituste analüüsimiseks ja nende omavaheliseks võrdlemiseks ning liigitamiseks kasvab pidevalt. Seda on vaja ürituste parema turundamise ja korraldamise jaoks. Erinevaid üritusi on võimalik võrrelda siis, kui on kasutatud ühtseid meetodeid andmete saamiseks. Praeguseks hetkeks ei ole loodud ühtset metoodikat ürituste omavaheliseks võrdlemiseks kvantitatiivsete näitajate abil. Ettevõtlike Arendamise Sihtasutusel (EAS) on olemas andmebaas üritustest, nimedest ja kuupäevadest, kuid puuduvad statistilised näitajad. Peale selle on olemas andmebaas rahastuse saanud ürituste taotlustest saadava infoga, kuid see pole täielik. EASil on huvi saada paremat infot Eestis toimuvate ürituste kohta, et teha paremaid investeeringuid ürituste toetusteks. Ürituste tuvastamise vastu on huvi ettevõttel OÜ Positium LBS, kes haldab mobiilpositsioneerimise andmeid ning soovib nendest andmetest eristada üritused. Automaatne ürituste tuvastamine annaks võimaluse kvantitatiivset statistikat koguda ürituste kohta ja jagada seda kolmandate osapooltega, kes sellest võiksid huvitatud olla.

Käesoleva töö eesmärgiks on koostada algoritm passiivse mobiilpositsioneerimise andmete abil, millega on võimalik:

- tuvastada üritusi,
- leida ürituse algus- ja lõpuaeg ning
- mõõta ürituste geograafilist ulatust.

Lisaks on eesmärgiks testida algoritmi ja arutleda algoritmi abil tuvastatud ürituste geograafilise ulatuse, algus- ja lõpuaja ning üritusel osalenud inimeste tunnuste kohta ning uurida, kuidas oleks võimalik tuvastatud üritusi liigitada. Ürituste tuvastamiseks loodi algoritm 2013. aasta kõigi eestlaste ja välismaalaste andmete põhjal ja analüüsiti selle tulemuste ajalis-ruumilist jaotumist.

Töö koosneb neljast peatükist. Teoreetilisest osast, kus käsitletakse ürituse definitsiooni, ürituste liigitamist ja varasemaid ürituste uurimusmeetodeid. Teine osa koosneb mobiilpositsioneerimise andmete, algoritmi ning analüüsi metoodika kirjeldustest. Kolmandas osas antakse ülevaade algoritmi kasutamise tulemustest. Viimasena arutatakse saadud tulemuste põhjal koostatud algoritmi täpsust ja selles tekkivaid vigu.

Töö käigus valminud algoritmi varalised õigused kuuluvad ettevõttele OÜ Positium LBS.

1. Teoreetilised lähtekohad

1.1 Ürituse definitsioon

Üritused on üks tugitalasid turismis ning omavad suurt osa arengu ja turunduse plaanides pea igas sihtkohas (Getz 2008). Üritused on olnud pikalt väga tähtsal kohal ühiskonnas. Teadaolevad esimesed olümpiamängud toimusid Vana-Kreekas juba 776. aastal eKr ning sellest ajast on toimunud pidevalt erineva sisu ja suurusega üritusi üle maailma (Jago, L., Shaw, R., 1998 cit. Nilbe 2011). Turismiüritusi on uuritud, et kirjeldada sihtkoha arengut ja turundusstrateegiat. Peamiseks eesmärgiks on paremini mõista ürituse kõiki majanduslikke külgi (Getz 2005).

Siiani on üritusi kirjeldatud sisu: planeerimise ja eesmärgi järgi, kuid ainult mobiilpositsioneerimise andmete pealt ei ole võimalik tuvastada ürituse kvalitatiivseid tunnuseid. Tuginedes Getzile (2005) on üritused ajaliselt esinevad, planeeritud ja planeerimata ning kindla ajalise pikkusega. Planeeritud üritusel on ajaline kestvus määratud ja avalikustatud. Üks põhimõte laieneb kõigile üritustele: nad on ajutised ja iga üritus on unikaalne oma korralduse, programmi, ülesehituse ning küllastajate poolest (Getz 2005). Seda põhimõtet toetab ka Carry definitsioon, kus üritus on piiratud kestvusega ühekordne või korduv sündmus, mis ei toimu sagedamini kui kord aastas. Üritusel on olemas programm, korraldav organisatsioon, osalejad ja on avalik kõigile. Peale sisulise eesmärgi (näiteks sport, kultuur, religioon) arendatakse üritusi eelkõige selleks, et korraldatava linna, piirkonna või riigi teadlikkust, köitvust ja kasumlikkust tõsta ning suurendada antud turismisihtkohta. Selle definitsiooni järgi on üritus näiteks olümpiamängud ning ka paljud festivalid ja laadad, samas juhatuse koosolekud ja regulaarsed kokkusaamised ei ole üritused, sest toimuvad pidevalt. Samuti ei ole teater, kino ja erinevate spordialade liigad üritused, sest need on korduvad ja regulaarsed. Kui kontserdid ja festivalid on avalikud üritused, siis koosolekute või konverentside jaoks on vajalik kutse. Seega pole need sündmused avalikud ja ei klassifitseeri üritusena (Carrey 1994 cit. Mossberg 2000).

Lähtudes definitsioonidest saab inimeste liikumisi jagada kaheks: regulaarseteks ja mitteregulaarseteks ehk ajutisteks. Püsivad käitumuslikud harjumused on kindla ööpäevase, nädalase ja aastase rütmiga. Püsivast rütmist erinevat käitumist võib pidada ajutiseks. Ajutisel käitumise erinevusel on olemas kindel algus ja lõpp, mille jooksul erineb käitumine regulaarsest rütmist. Ürituse küllastajad moodustavad iga ürituse jaoks unikaalse kogumi, mis on ürituse siseselt sarnasem kui teiste üritustega. Sellest lähtudes sobib

mobiilpositsioneerimise andmete jaoks definitsioon: üritus on inimeste tavapärasest erinev käitumine, millel on kindel algus ja lõpp ning mida ühendab ainulaadne külastajate kogum. Autorile teadaolevalt ürituste geograafilise ulatuse kohta varasem käsitlus puudub.

Vaadates mobiilpositsioneerimise andmeid teatud kindlas ajas ja ruumi klastris on võimalik hinnata selles klastris olevate mobiilide hulka. See annab meile võimaluse teada, millised inimesed võiksid olla mõne ürituse külastajad. Mobiilpositsioneerimise andmed võimaldavad anda ülevaate ürituse objektiivsete ja kvantitatiivsete näitajate kaudu, näiteks hinnatav kohalike ja välismaalaste vahekord, oletatav külastajate arv, ürituse geograafiline ja ajaline ulatus.

1.2 Ürituste uuringute lähteandmed

Inimeste liikumismustrite uurimisel on üheks fookuseks turistid. Peamiseks põhjuseks on turismi parem mõistmine ja seeläbi ka parem arusaam turistide liikumisest. Turismi mõistmine annab teadmisi turismi paremaks planeerimiseks ning majandusliku ja sotsiaalse mõju suurendamiseks. Turistide mobiilsust on uuritud varasemalt küsitluste, vaatluste või intervjuudega. Need on tömahukad, subjektiivsed ja vähese usaldusväärsusega. Tehnoloogia arenguga on kasutusele võetud mitmeid objektiivseid tehnoloogiaid. Üheks neist on GNSS (globaalne satelliitnavigatsioonisüsteem), mille abil logimisseadmed määravad oma asukoha täpsusega kuni 1m, mis võrreldes teiste meetoditega on tunduvalt täpsem ja objektiivsem. Sellel meetodil on omad puudujäägid: asukohta ei saa määrata siseruumides ja piisava hulga andmete kätte saamine on raskendatud (Shoval et al 2011). Lisaks teeb andmete kogumise problemaatiliseks seadmete lühikene akukestvus ning sagedasti unustatakse seade endaga kaasa võtmast. GNSS-i on võimalik kasutada ka mobiiltelefonides. See eeldab kasutaja nõusolekut ja vastava mobiilirakenduse paigaldamist. Mobiiltelefonide akukestvus on veelgi väiksem võrreldes GNSS seadmetega. Mobiiltelefoni eeliseks on selle pidev kaasaskandmine endaga ning see on igapäevane isiklik. Kasutatakse ka sinihamba tehnoloogiat määramaks mööduvaid seadmeid. Kuid statsionaarsed logimisseadmed puuduvad (Versichele et al 2014). Samal põhimõttel kasutatakse ka raadiosagedustuvastust, mille levik on väiksem, kuid mis ei nõua toidet jälgitava poolt. Antud meetodi puudusteks on piiratud raadius, mille piires on võimalik tuvastada kiipe. Peamiselt kasutatakse selliseid süsteeme haiglates ja tööstustes, kus on ruumidesse sissepääs tagatud raadiosagedustuvastusega (Öztaysi et al 2009). Sarnasel

kaugtuvastuse meetodil on proovitud WIFI seadmete abil tuvastada ja uurida inimeste mobiilsust ühe massiürituste raames. See ei nõua osalejate poolset sekkumist (Bonné et al 2013). Antud meetodil pole võimalik tuvastada isikuid ning andmete kogumisel on tagatud anonüümsus.

Eelnevad meetodid on sobivad ühe ürituse raames inimeste liikuvuse tuvastamiseks, kuid aastaringselt ja üleriigiliseks katmiseks on antud meetodid andmete kogumiseks kordades töömahukamad ja kulukamad kui mobiilpositsioneerimise abil kogutud andmed. Mobiilpositsioneerimise andmeid on kasutatud varem turistide regionaalsete liikumismustrite uurimiseks (Ahas et al 2007) ning turismiürituste tagamaade tuvastamiseks (Nilbe, 2011). Päringuid ürituste kohta mobiilpositsioneerimise andmetelt on tehtud varemgi, kuid geograafiline ala on ära määratud käsitsi mastide järgi ning ajaline kestvus on võetud kuupäeva täpsusega (Nilbe 2008). Autorile teadaolevalt ei ole varem mobiilpositsioneerimis- andmeilt ürituste tuvastamist automatiseeritud.

1.3 Ürituste liigitamine

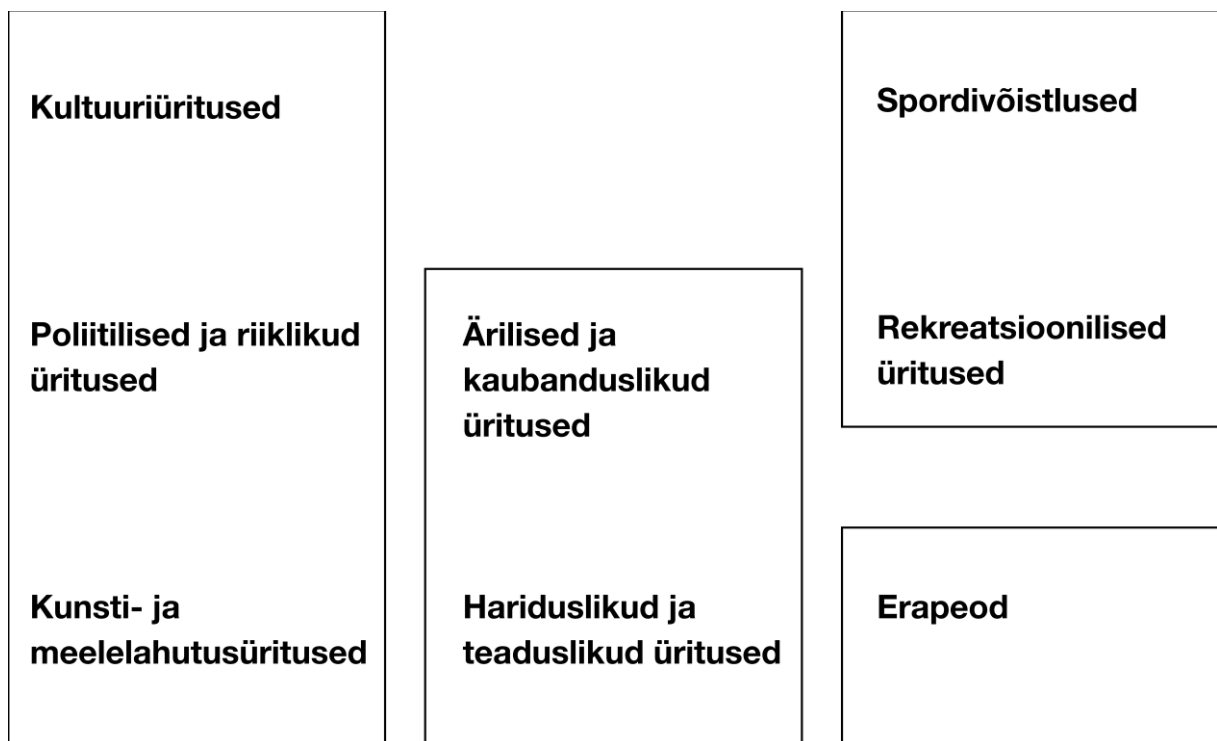
Üritusi korraldatakse väga erineva sisu ja suurusega. Nende paremaks mõistmiseks on üritusi hakatud jaotama sisu ja suuruse alusel erinevateks tüüpideks (Getz 2005; Jago, Shaw 1998 cit. Nilbe 2011). Jagades üritused sisu ja suuruse järgi, on võimalik paremini võrrelda sama liiki üritusi ning leida antud üritusele iseloomulike tunnuseid.

Üritusi saab jaotada sisu järgi mitut moodi. Nii palju kui on autoreid on ka erinevaid arvamusi. Seega pole väljakujunenud ühte kindlat teooriat. Sellepärast sõltub uuringu eesmärgist, millist klassifikatsiooni kasutada (Nilbe 2011). Ürituse liik määratakse ainult planeeritud üritustel, sest neil on olemas kindel eesmärk ning korraldaja, mille järgi liigitada.

Üks põhjalikumaid käsitlusi on loonud Getz (2005;2008), kus ta proovib leida peamised kategooriad eristamaks avalikke üritusi väikestest ja eraüritustest. Autori arvates pole selline jaotus kindlasti täielik ning olenevalt ürituse eesmärgist ja olukorrast võib üritusi jaotuda mitmesse kategooriasse samaaegselt. Kõiki antud kategooriate üritusi on võimalik leida igast kultuurist ja kogukonnast. Iga kategooria jaguneb omakorda ürituste klassideks.

- Kultuuriüritused
 - Mälestusteenistused
 - Karnevalid
 - Festivalid
 - Religioossed üritused
 - Paraadid
- Poliitilised ja riiklikud üritused
 - Väga tähtsate isikute (VIP) külastused
 - Tippkohtumised
 - Poliitilised üritused
- Kunsti- ja meelelahutusüritused
 - Kontserdid
 - Auhinnatseremooniad
 - Näitused
- Ärilised ja kaubanduslikud üritused
 - Kohtumised
 - Konverentsid
 - Reklaamid
 - Raha kogumise üritused
 - Messid
 - Laadad
 - Turgude motivatsiooniüritused
- Haridus- ja teaduslikud üritused
 - Konverentsid
 - Kongressid
 - Seminarid
 - Töötoad
 - Lühikursused
- Spordivõistlused
 - Amatööridele ja professionaalide
 - Pealtvaatajatele ja osalejatele mõeldud üritused
- Rekreatsioonilised üritused
 - Meelelahutuslikud spordimängud

- Erapeod
 - Isiklikud üritused
 - Tähtpäevad
 - Ärasaatmised
 - Perepuhkused
 - Ühiskondlikud üritused
 - Kokkutulekud
 - Galad
 - Pulmad



Joonis 1. Ürituste liikide jaotumine sisu järgi (Getz 2005).

Ürituse suurust on võimalik mõõta füüsiliste (toimumisala, inimeste arvu, tagamaasuuruse), majanduslike (käibe, tulu) ja sotsiaal-psühholoogiliste (imago, reklaam) näitajatega. Nilbe (2011) on jaotanud üritused külastajate arvu järgi kolmeks:

- väikesemõõtmelised üritused;
- keskmise suurusega üritused;
- suurüritused.

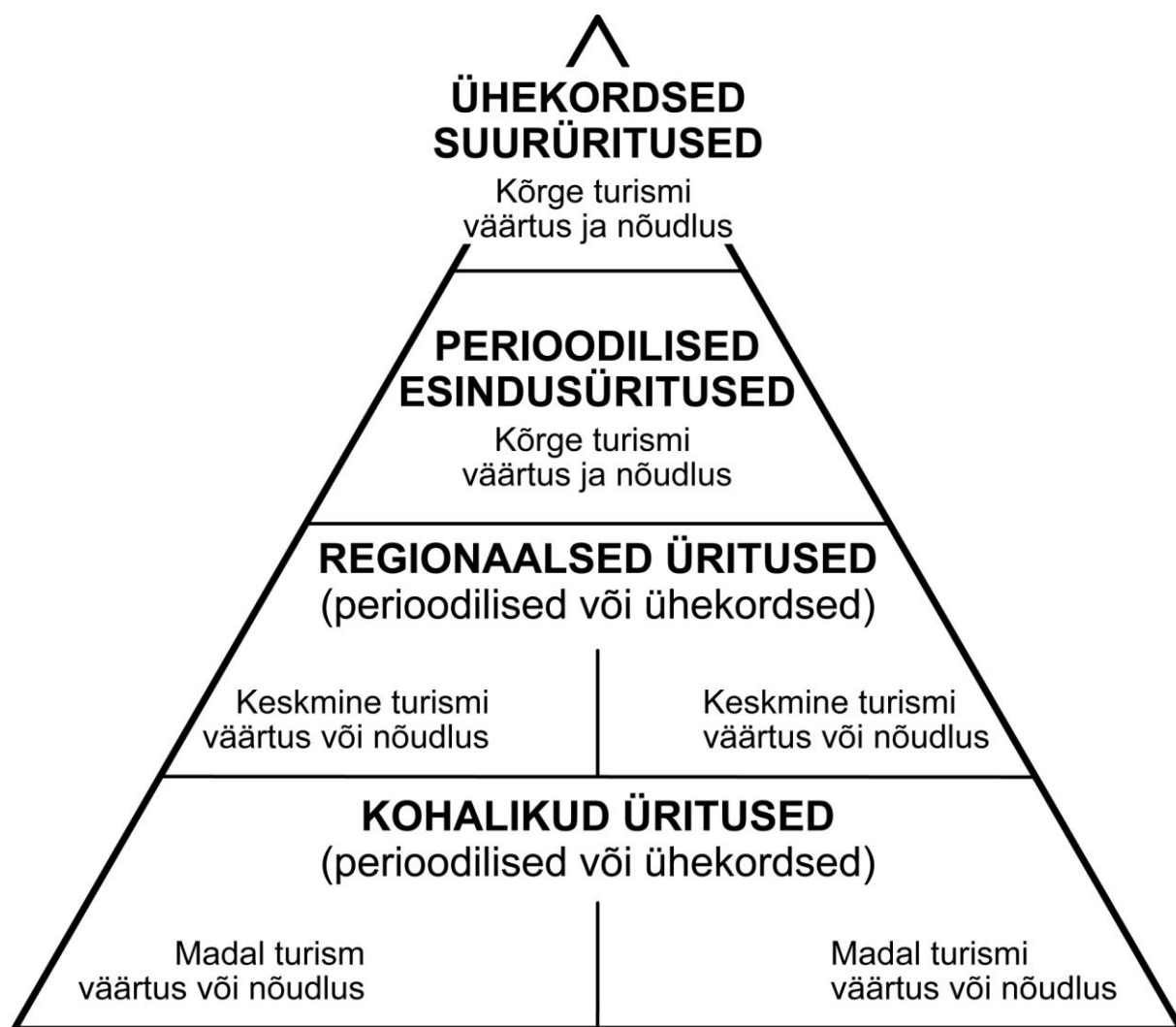
Väikesemõõtmelised üritused on mõeldud kogukonnale ja need on näiteks tänaväitused 2000 külastajaga. Keskmise suurusega üritused on regionaalsed ja selle alla kuuluvad festivalid 50 000 külastajaga. Suuremõõtmelised üritused omavad rahvusvahelist tähtsust. Selleks võivad olla ühekordsed suurüritused, näiteks Eurovisiooni korraldamine. Ürituse suurenedes suureneb ka lähiriikide külastajate osatähtsus osalejatest (Nilbe 2011).

Igal kogukonnal ja regioonil on olemas portfoolio üritustest, mida saab jagada tüübi, suuruse, hooaja ja mõju järgi. Selleks, et saavutada turismi eesmärgi paremini oleks vaja mitmekülgset ja universaalset ürituste portfoolio hindamist. Getz (2005) on loonud funktsionaalsusel loodud tüpoloogia. See põhineb kahel kriteeriumil:

- nõudlus, mida mõõdetakse meelitatud turistide arvu trendide järgi;
- väärtus, mis sisaldab endas teiste turismi eesmärkide täitmist. See võib mõõta kasvu potentsiaali, turuosa suuruse, kvaliteeti, üldpildi parandamise, kogukonna toetuse, miljööväärtuse, majandusliku kasu, jätkusuutlikkuse või sobivuse alusel.

Antud jaotuse alusel jagunevad üritused 4 gruppi:

- ühekordsed suurüritused, mis on tavaliselt ülemaailmsed üritused, mille korraldamiseks on vaja võita ühekordse korraldamise õigus, näiteks olümpiamängud;
- korduvad esindusüritused, on seotud kindla kogukonnaga ja ei saa eksisteerida sellest eraldi;
- regionaalsed üritused, on tugev side kohaliku kogukonnaga kuid lisaks kohalikele külastajatele meelitab kohale ka turiste ning võivad olla nii ühekordsed kui korduvad;
- kohalikud üritused, mis on suunatud kohalikule kogukonnale ning mille eesmärgiks on piirkonna või kogukonna nähtavamaks muutmine, tugevdada kogukonna ühtekuuluvustunnet ja uute ideede, kogemuste tekitamine.



Joonis 2. Funktsionaalsusel põhinev ürituste jaotumine (Getz 2005).

2. Andmed ja metoodika

Selles peatükis tutvustakse töös kasutatud andmeid, nende kogumise meetodeid ning andmetöötluse printsiipe. Lisaks antakse ülevaade turismiürituste tuvastamise metoodikast.

Andmete töötlemiseks ja hoidmiseks kasutati PostgreSQL andmebaasi koos PostGIS laiendiga. PostGIS lisab PostgreSQL andmebaasihaldurile võimaluse hallata ja töödelda ruumilisi andmeid. Lisaks kasutati töös andmete töötlemiseks Pythoni programmeerimiskeelt ja MS Excel tabelarvutustarkvara. Visualiseerimiseks kasutati QGIS geoinfotarkvara ja kujundusprogrammi Adobe Illustrator.

2.1 Mobiilpositsioneerimise andmed

Antud töös on kasutatud isikustamata passiivseid mobiilpositsioneerimise andmeid ühelt Eesti mobiilioperaatorilt, mis on kogutud ajavahemikul 1. jaanuar 2013 kuni 31. detsember 2013 ettevõtte Positium LBS poolt. Algandmed on isikustamata, see tähendab, et igale kasutajale on loodud unikaalne, isikust ja SIM-kaardist sõltumatu tunnus. See tagab inimeste privaatsuse ja on kooskõlas Eesti ja Euroopa andmekaitse seadustega (Ahas et al 2007).

Iga kõnetoimingu jaoks oli teada unikaalne ID, kõnetoimingu aeg, antenni ID. Lisaks oli teada SIM-kaardi päritolumaa ja antenni ligikaudsed koordinaadid. Kõnetoimingu alla kuuluvad väljuv kõne, SMS, MMS ja teised operaatori poolt maksustatavad kõnetoimingute teenused.

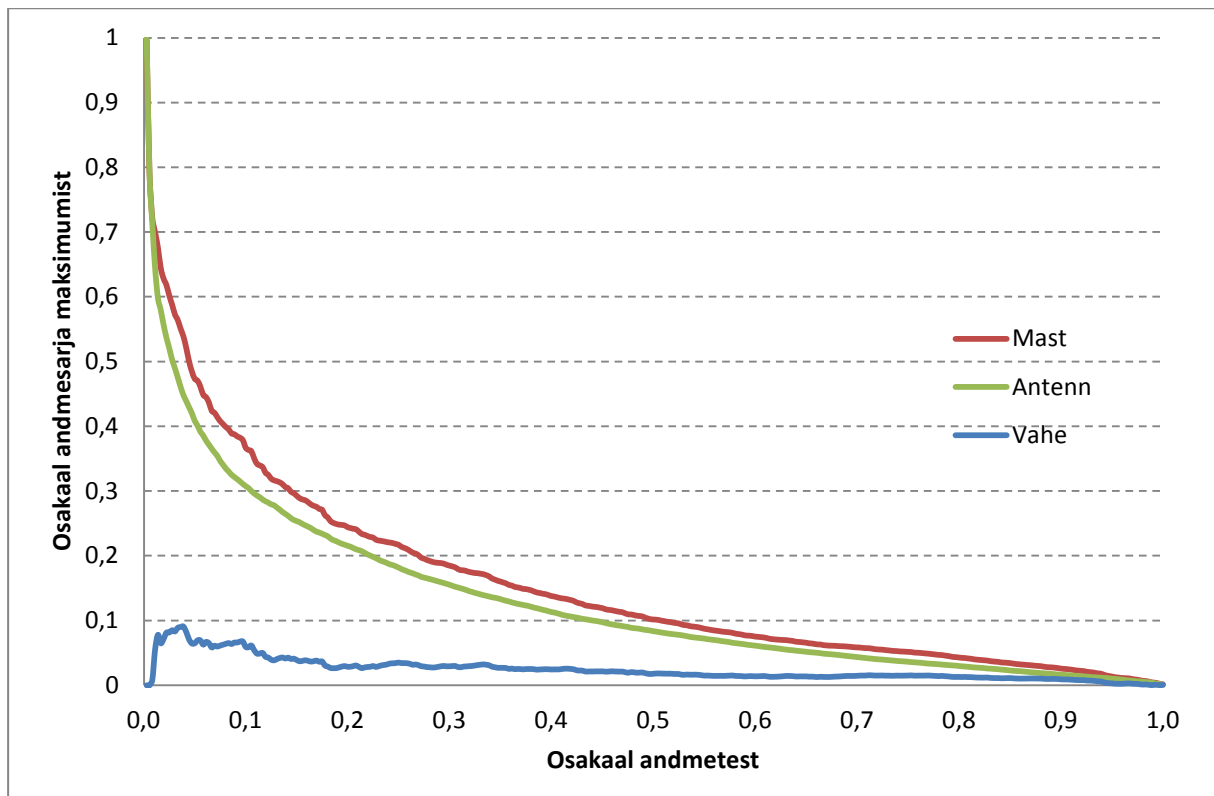
Mobiiltelefon ühendab ennast enamasti kõige lähema antenniga. Mobiilmastide võrgustik pole ühtlane vaid jälgib rahvastiku tihedust ja transporditaristut (Ahas et al 2010). Linnades on üks mast 200 m² kohta ja maapiirkondades võib see olla kuni 100 km² (Ahas et al 2007). See näitab ka mobiiltelefoni positsioneerimise täpsust, mis on otseses sõltuvuses mastide tihedusest.

Mobiilmastide võrgustik pole püsiv ning on pidevalt muutuv. Peale selle teeb andmete töötlemise keerukaks erinevate sidetehnoloogiate kasutamine mobiilioperaatori võrgus ja nende omavaheline sidumine ühtseks tervikuks (Ahas et al 2010).

2.2 Andmetöötlus

Algandmetest tehti koondtabel, mida on lihtsam töödelda. Koostatud tabeli andmed on grupeeritud masti ja tunni kaupa nõ mast-tundideks. Iga masti ja iga tunni jaoks oli

andmebaasis eraldi rida. Kõned on grupeeritud masti järgi, sest nii on kõned ühtlasemalt jaotunud kui antennide järgi (joonis 3).

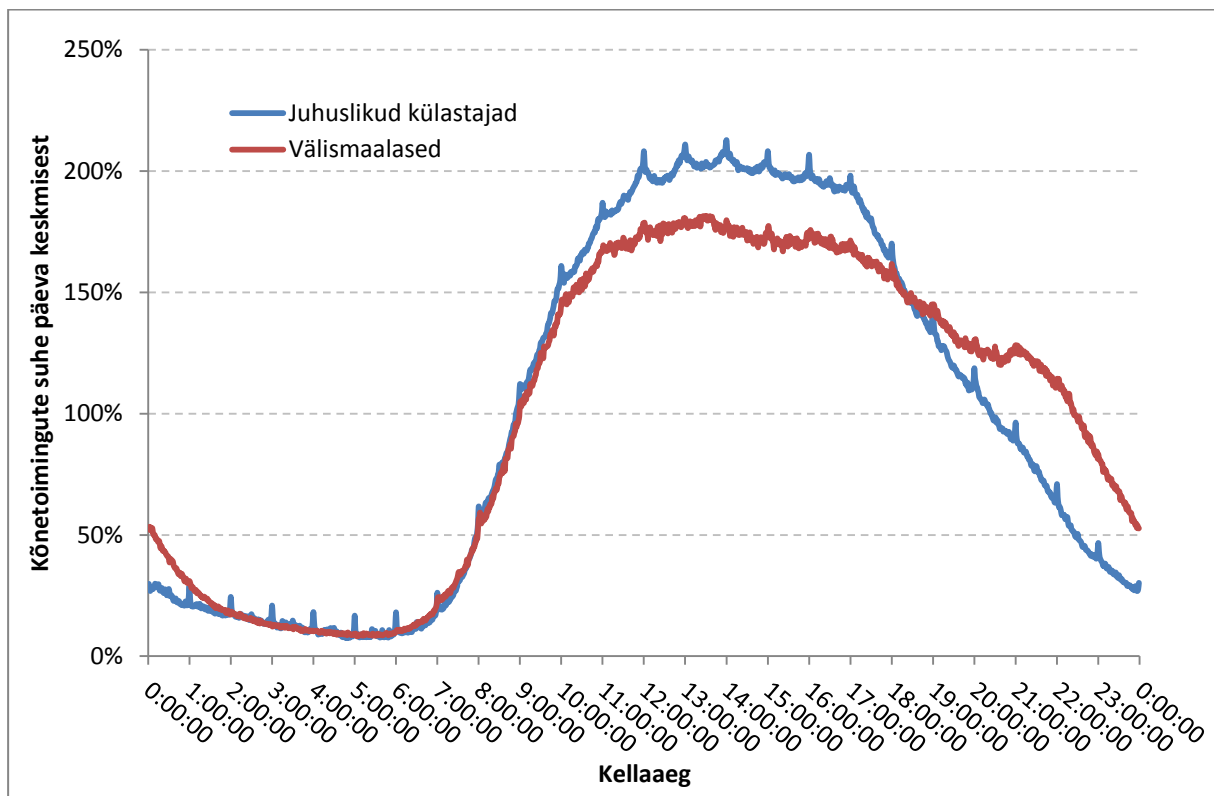


Joonis 3. Masti ja antenni vahel kõnetoimingute jagunemine.

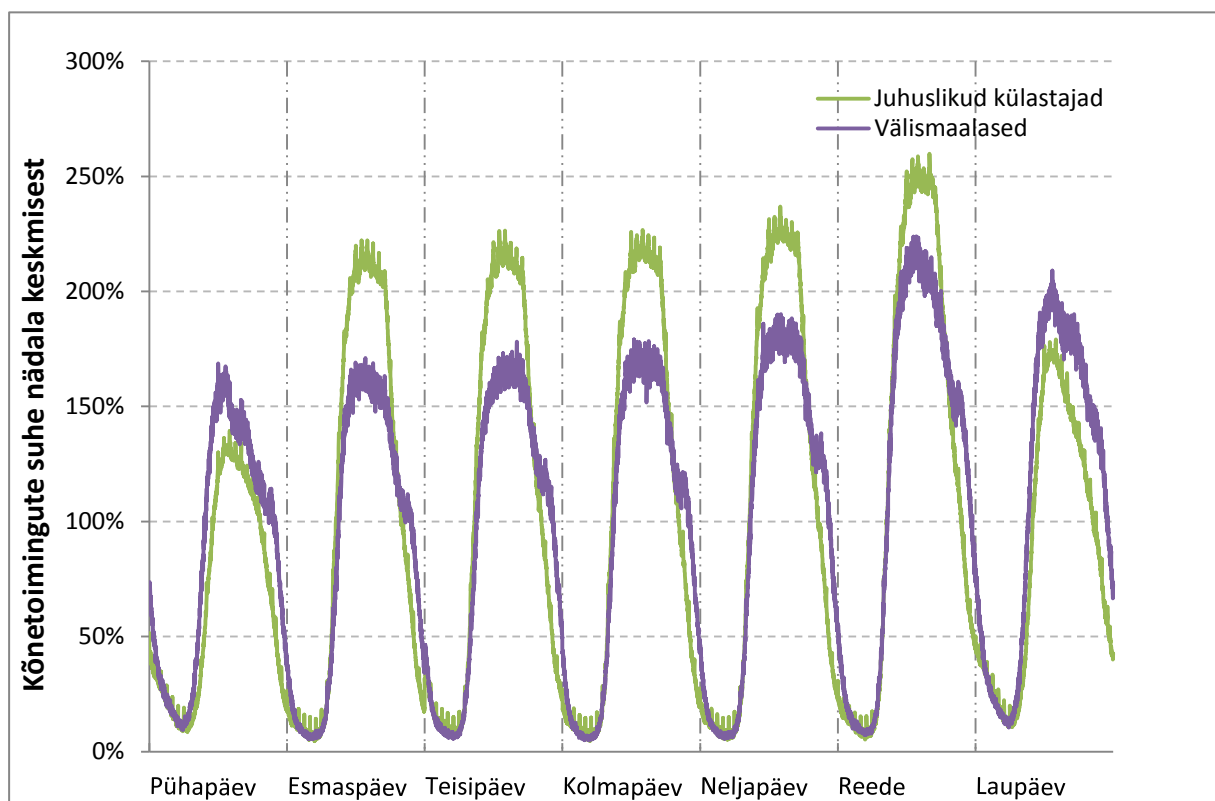
Nii masti kui ka antenni andmete kohta arvutati variatsioonikoeffitsient, mis näitab aritmeetilise keskmise järgi standardiseeritud suhtelist hajuvust. Masti järgi on variatsioonikoeffitsient 98,4% ja antenni järgi 105,6%. Kõned jaotati masti järgi, sest see annab piisava täpsuse ürituse asukoha määramiseks. Samas on kõnede arv mastide vahel ühtlasemalt jaotunud kui antennide vahel. Kõned jaotati tunni kaupa, sest nii jaotus igasse tundi statistiliselt kriitiline arv kõnesid, et see oleks piisav täpsus ürituse kestvuse tuvastamiseks. Iga tunni ja masti kohta summeeriti kokku kogu kõnede arv, unikaalsed välismaa ja Eestis kasutajad, kes antud mast-tunni jooksul tegid vähemalt ühe kõnetoimingu. Eesti kasutajad jagati kaheks: ankurpunktiga, kelle üks ankurpunktidest on antud masti juures ja juhuslikud, kes on teinud antud tunnil kõnetoimingu, kuid ei oma antud masti juures ankurpunkti. Ankurpunktid arvutati Tartu Ülikooli geograafia osakonnas valminud algoritmi järgi. Kõigile Eesti kõneoperaatori kasutajatele arvutati ankurpunktid, kus inimene tegi ühe kuu jooksul vähemalt seitsmel erineval päeval kõnetoiminguid (Ahas et al 2010). Selline meetod annab võimaluse jagada kõnesid regulaarseteks, kõne toimus mõne kasutajale

täendusliku ankurpunkti lähedal, ja tavapärasest erinevaks ehk juhuslikuks. Kuna eelnevalt on defineeritud üritust kui tavapärasest erinevat sündmus, siis kasutame neid kõnesid, mis on tehtud tavapärasest erinevas kohas. Nii saamegi rääkida väliturismist ja juhuslikest külastustest, kus välituristik on mobiilinumbriga omanik mõnest välisriigist, kes kasutab Eestis rändlusteenust (*roaming*) ja juhuslikuks külastajaks on mobiilinumbriga omanik Eesti mobiilivõrgus, kes külastab neid kohti harva (ei ole igapäevase tegevusruumi osa).

Välismaalaste ning juhuslikel külastajate tehtud kõnede arvul ja unikaalsetel kasutajatel on kindel ööpäevane, nädalane ja aastane rütm (joonised 4 ja 5) juhuslikud külastajad teevad rohkem kõnesid tööajal, alates hommikul kella üheksast kuni kella kuueni õhtul. Juhuslikel külastajatel kasvab igal täistunnil kõnetoimingute arv. See on seotud täistunnil algavate kohtumistega. Välismaalastel on kõnetoimingute aktiivsus õhtul kella kuuest kuni öösel kella kaheni kõrgem kui eestlastel. See on tingitud vabamast ajagraafikust kui juhuslikel külastajatel, kellel on rohkem kohustusi. Peale kella kahte öösel nii juhuslikel külastajate kui ka välismaalaste kõnetoimingute arvu suhe mediaani võrdsustub.



Joonis 4. Päeva rütm standardiseeritud keskmise järgi.

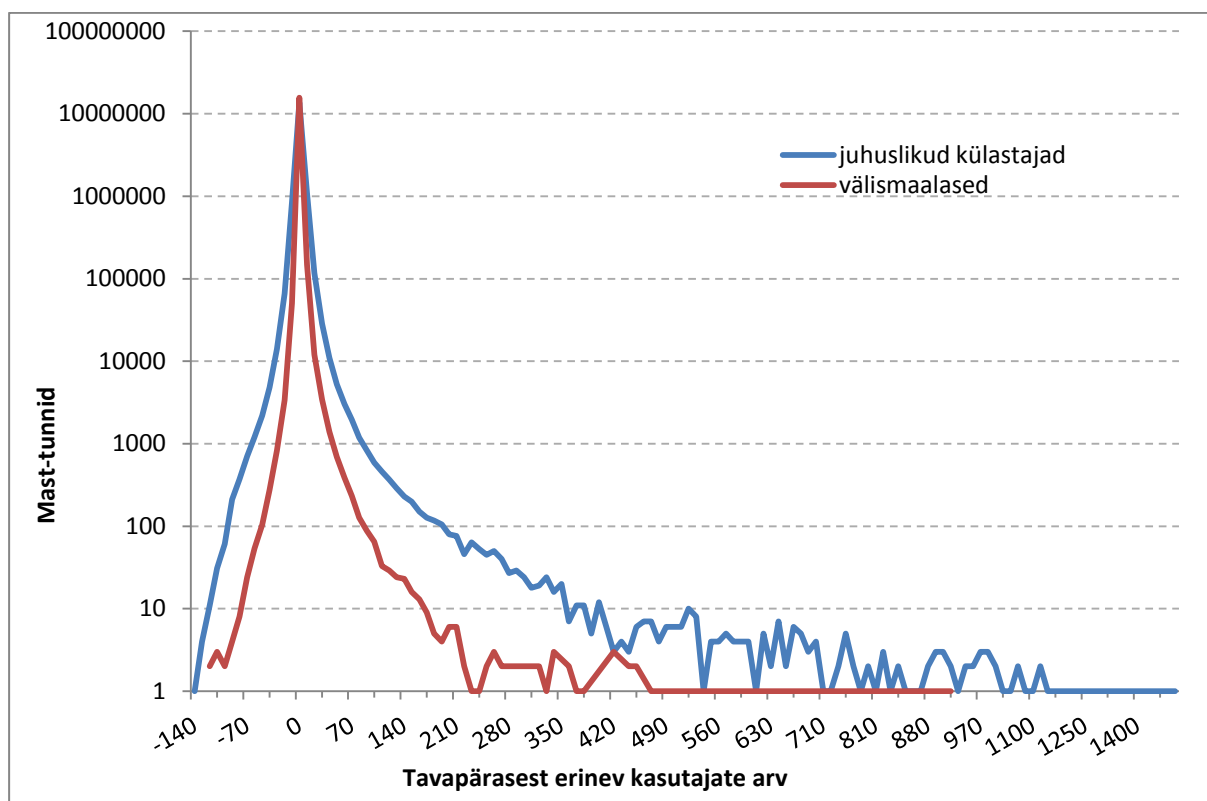


Joonis 5. Nädala rütm standardiseeritud keskmise järgi.

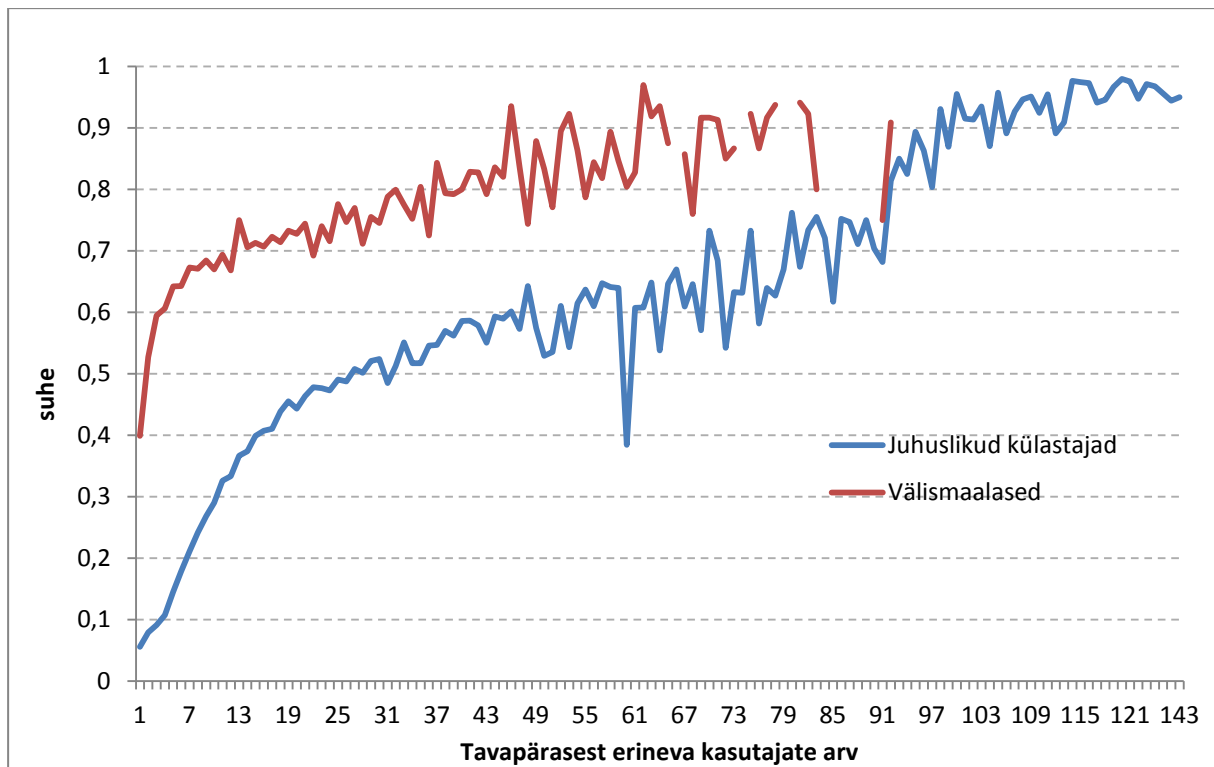
Paremaks ürituste tuvastamiseks tuleks leida antud tunnil ja masti juures tavapärasest erinev arv kasutajaid. Selleks, et leida iga päeva ja masti jaoks tavapärane juhuslike külastajate ja välismaalaste kasutajate arv, leiti iga tunni ning masti jaoks viie eelneva ja järgneva sama nädalapäeva ja kellaaja mediaan. See väljendab antud ajahetke ja masti tavapärast kasutajate arvu, eeldades, et antud ajahetkel nädalas on eelneva ja järgneva viie nädala jooksul olnud kuni viis üritust. Suvisel turismi hooajal võib-olla üritusi mitmel järjestikusel nädalal järjest samal ajahetkel nädalas. Kuna ka aasta lõikes on kindel rütm, siis ei saa võtta ajaperioodiks tervet aastat.

Peale seda leiti antud tunni tavapärasest erinev kasutajate arv nii juhuslikele kui ka välismaalaste külastajatele. Selle põhjal arvutati iga mast-tunni jaoks tavapärasest erineva kasutajate arv ehk siis leiti viie eelneva ja järgneva nädala mediaani suhe. Kuna juhuslikel külastajatel ja välismaalaste kasutajatel on erinev päeva, nädala ja aasta rütm, siis arvutati mõlema jaoks eraldi tavapärasest erinev kasutajate arv antud mast-tunnil. Nii juhuslikel külastajatel kui ka välismaalaste tavapärasest erinev kasutajate arvu mediaan on 0. Mediaani järgi võib öelda, et tavapärane kasutajate arv on arvutatud täpselt. Sagedusjooniselt on näha

(joonis 6), et saba on positiivsete väärtuste suunas, sest üritustel on kasutajate arv tavapärasest suurem. Võib eeldada, et kui üritusi ei toimuks, oleks sagedus jaotunud sümmeetriliselt mediaanist. See annab võimaluse vaadata kui suure osa moodustab tavapärasest suurema ja väiksema kasutajate arvu ehk ürituste mast-tundide arv vastava tavapärasest erineva kasutajate arvu korral mast-tundide arvust. Selleks võeti aluseks negatiivsed tavapärasest erineva kasutajate arvu väärtused ning vaadati kui suure osa moodustavad negatiivsed tavapärasest erinevate kasutajate arvu mast-tundide arv ja tema absoluutväärtuse tavapärasest erinevate kasutajate arvu mast-tundide arv vahe kõigist antud mast-tundide arvust (joonis 7). Selle põhjal leiti tavapärasest erinev kasutajate arv, kus eeldatavate ürituste mast-tundide arvu osakaal oli vähemalt 0,6 kõikidest antud tavapärasest erinevate kasutajate arvust. Leitud kasutajate arvust suuremad tavapärasest erinev kasutajate arvuga mast-tunnid valiti välja kui potentsiaalsed üritused. Kuna tavapärasest erineva kasutajate sagedusjaotus on erinev juhuslikel külastajatel ja välismaalastel, siis leiti mõlema jaoks eraldi tavapärasest erinev kasutajate arv, millest alates arvati potentsiaalsete ürituste hulka.

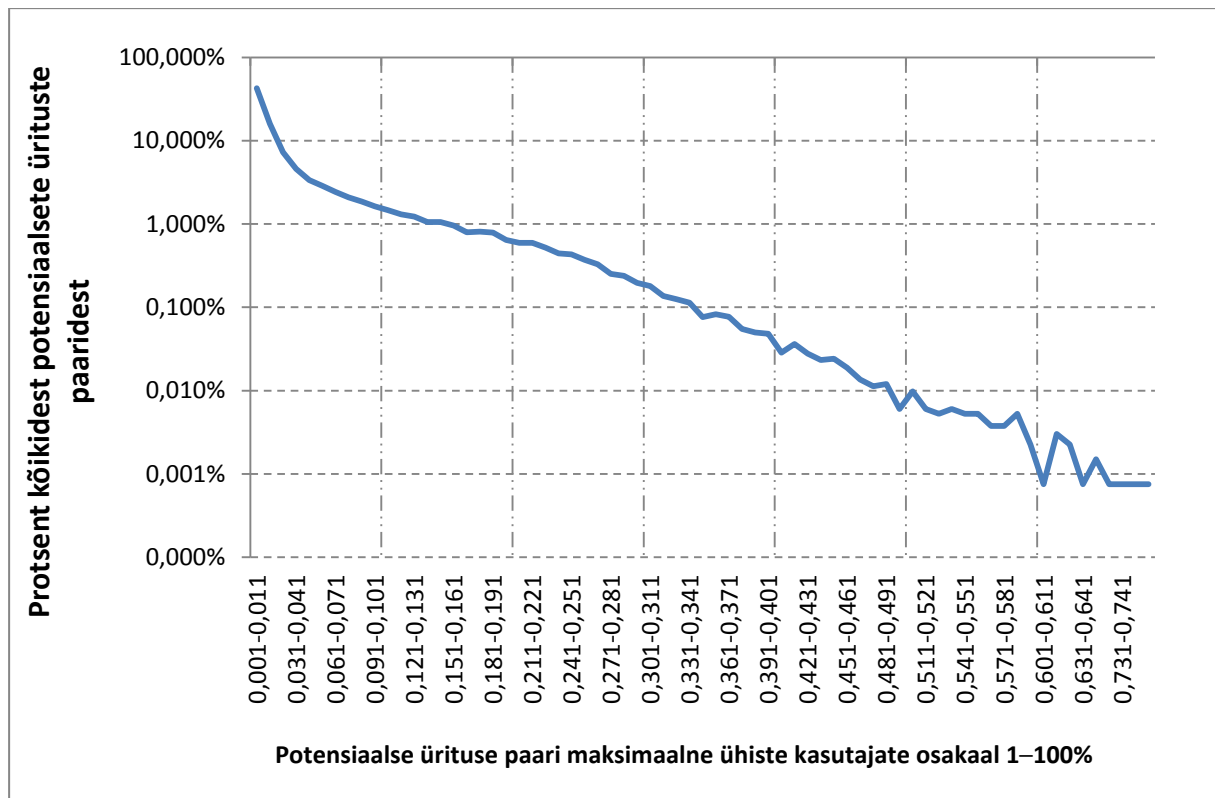


Joonis 6. Tavapärasest erineva kasutajatearvu sagedusjaotus.



Joonis 7. Tavapärasest erineva positiivse ja negatiivse kasutajate arvu vahe suhe tavapärasest erineva positiivsete kasutajatesse jaotus.

Välja valitud potentsiaalsed ürituse tunnid ja mastid grupeeriti DBSCAN (*Density-based spatial clustering of applications with noise*) ehk tihedusel põhinev ruumilise grupeerimise algoritmi abil. DBSCAN on tihedusel põhinev grupeerimine, kus sarnasuse teguriks valiti kahel potentsiaalsel üritusel ühised kasutajad ühe ööpäeva jooksul. Mida rohkem oli ühiseid kasutajaid üritustel, seda sarnasemad olid üritused. Potentsiaalsete ürituste ühisosa jaotusest (joonis 8) on välja jäetud ühisosa mitte omavad paarid. Epsiloniks valiti osakaal 0,1 ehk 10%. Joonisel 8 kõik seosed alates sarnasuse osakaalu vahemikust 0,091-0,101 võeti algoritmis kasutusele. minimaalseks grupisuuruseks 2. See tähendab, et kui ühel kahest potentsiaalsel ürituse mast-tunnil oli võrreldes teisega ühiste kasutajate arv kõigist kasutajatest vähemalt 10%, siis arvati need kaks potentsiaalset üritust üheks ürituseks. Minimaalne grupi suurus eraldab müra ja määrab ära kui suured peavad olema ürituse grupid selleks, et need moodustaksid uue ürituse grupi. Iga ürituste grupi jaoks arvutati tema geograafiline ulatus, ajaline kestvus, algus- ja lõppaeg, unikaalsed eesti ja välismaa kasutajad ning kõnede arv. Geograafiline ulatus arvutati iga ürituse mastide jaoks koostatud voronoide pindalade summana. Ürituse algusaeg võeti kõige varajasema ja lõppaeg kõige hilisema mast-tunni aja järgi.



Joonis 8. Potentsiaalsete ürituste paaride maksimaalne ühiste kasutajate jaotus.

2.3 Ürituste liigitamine

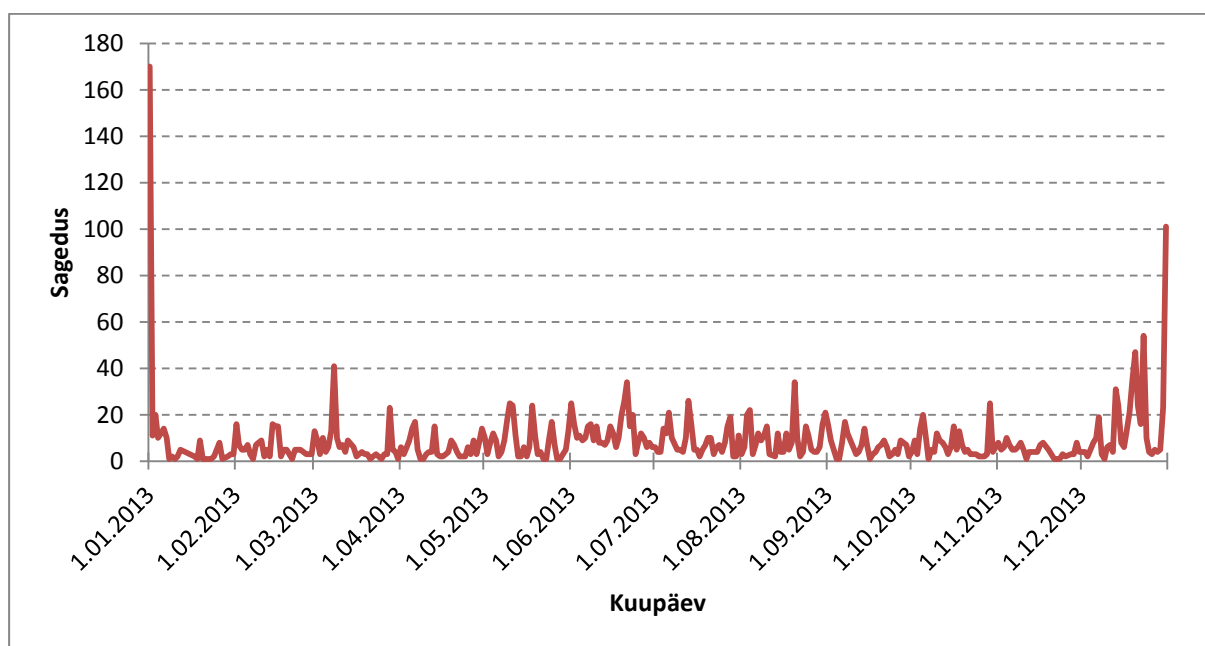
50-le kõige suuremale unikaalsete kasutajatega tuvastatud üritustele prooviti leida meedia ja interneti abil toimunud üritus. Kui üritus suudeti tuvastada siis määrati talle liik. Ürituste liigitamisel kasutati Getz (2005) loodud ürituste liigitamist. Kui üritust ei tuvastatud, siis liigitati see planeerimata ürituseks.

3. Tulemused

3.1 Automaatselt tuvastatud üritused

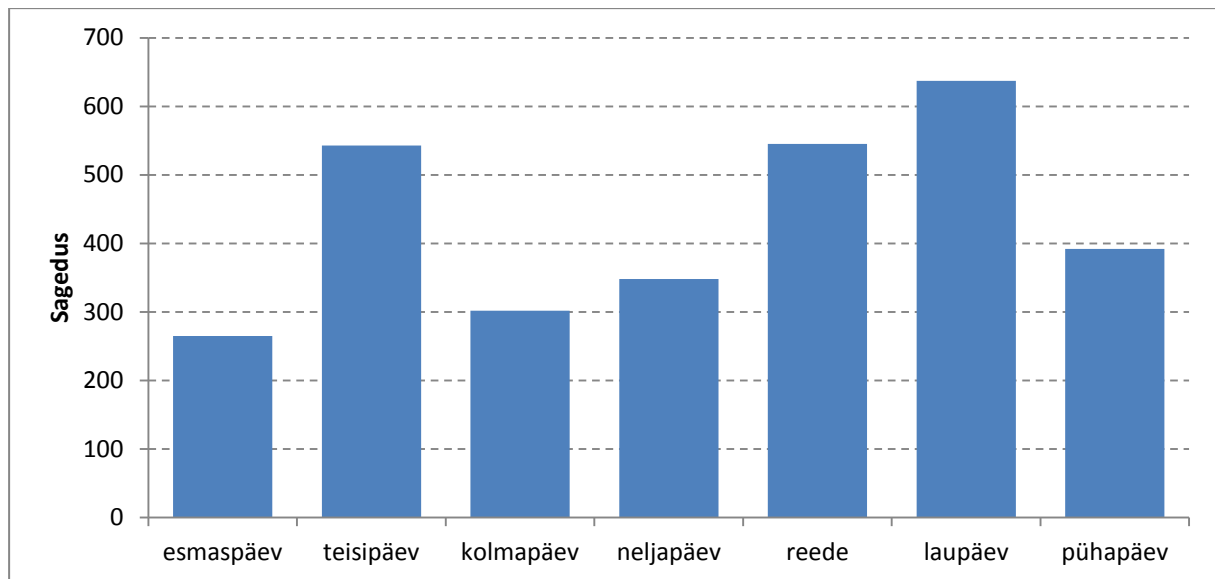
2013. aasta kõigist Eesti andmetest tuvastati 3032 üritust. Kokku oli kõigi ürituste kestvuseks 10245 mast-tundi, see teeb 0,11% kõikidest võimalikest mast-tundidest Eestis aasta jooksul.

Suurimateks üritusteks unikaalsete kasutajate alusel olid 20. augustil toimunud Robbie Williamsi kontsert Lauluväljakul ja uue aasta vastuvõtmine 1. jaanuari esimestel tundidel Tallinna kesklinnas. Vaadates ürituste jaotumist aasta lõikes (joonis 9) tulevad selgelt välja uue aasta vastuvõtt ja vana-aasta ärasaatmine. Kevadel tuleb esile ka 8. märts, mis on rahvusvaheline naistepäev. Suvel jaanipäev ja taasiseseisvumispäev 20. augustil. Sügisel eristub 29. oktoobri torm ja aasta lõpus jõulupühad. Juunist augusti lõpuni on üritusi enam kui teistel kuudel, neil kuudel tuvastati üle 1100 ürituse. Kõige vähem üritusi on novembris, kus tuvastati 329 üritust. Kõige enam 1395 üritust tuvastati detsembris.



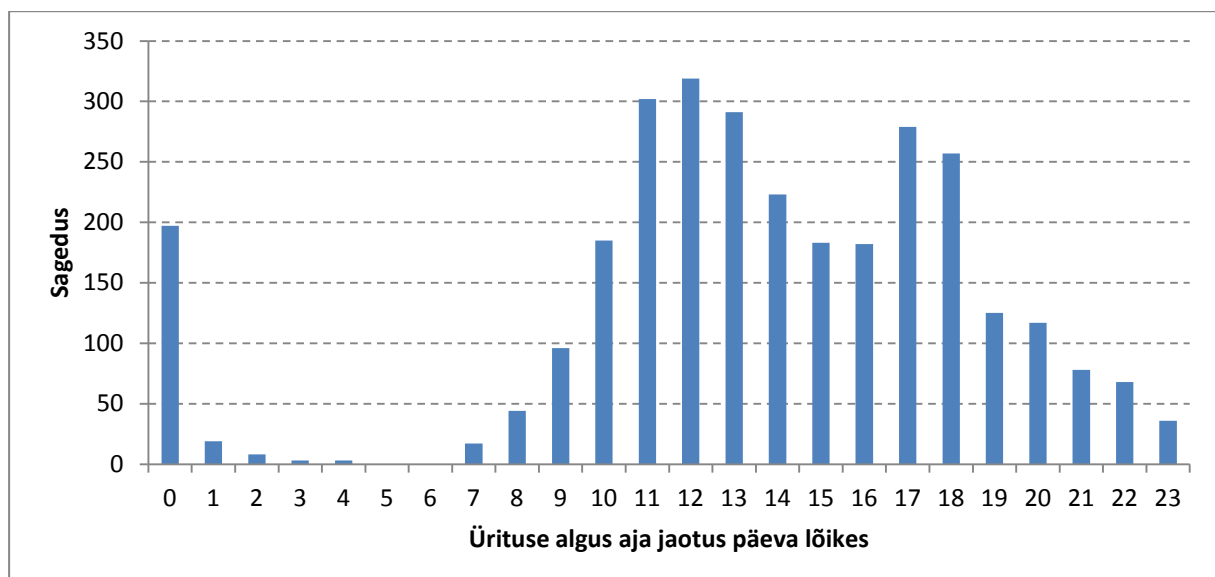
Joonis 9. Tuvastatud ürituste jagunemine aasta lõikes.

Ürituste jaotumine nädalapäevade lõikes (joonis 10) on mõjutatud aastavahetuse sattumisest kindlale nädalapäevale. Kuna 2013. aasta esimene ja viimane päev oli teisipäev, siis on ülejäänud aasta lõikes ürituste arv teisipäeval sarnane esmaspäeva või kolmapäevaga. Kui aastavahetuse üritused välja jätta, on enim üritusi nädalas reedel ja laupäeval.



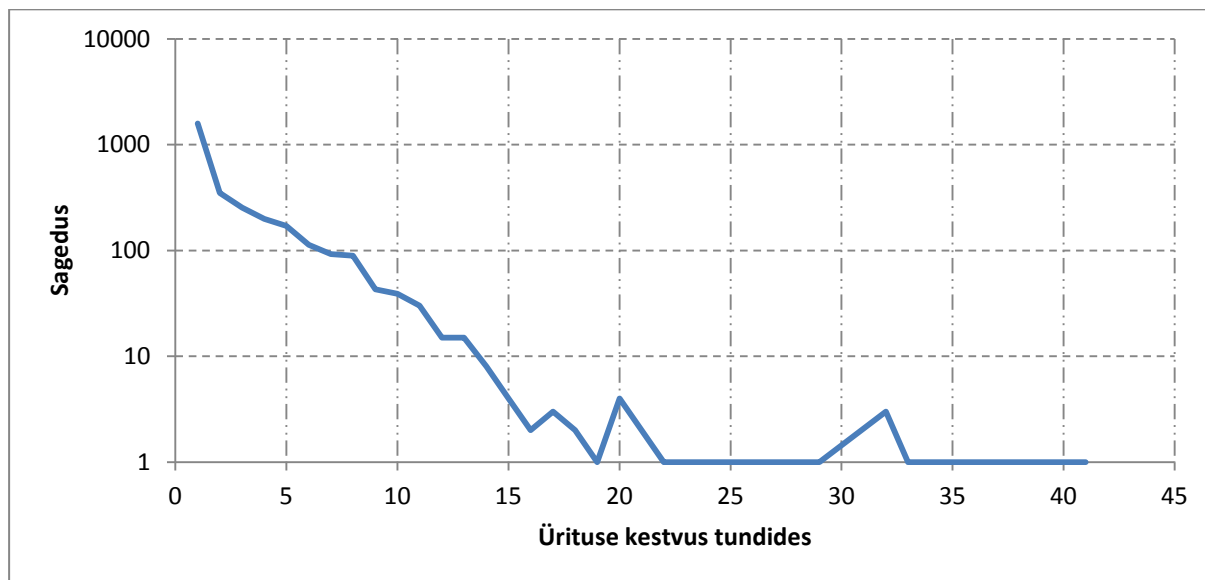
Joonis 10. Tuvastatud ürituste jaotumine nädalapäevade vahel.

Ööpäevases rütmis (joonis 11) eristub ürituste algusaja kaks tippu: lõuna - kell 12, ja õhtul - kella 17. Mõlemad tipud on seotud inimeste aktiivsusega ja liikumisega. Õhtune tipp on seotud tööpäeva lõppemisega. Lisaks on päeva esimesel tunnil tipp, mis on põhjustatud aastavahetusel üritustest üle Eesti. Varahommikul kella 5 ja 6 ajal ei tuvastatud ühtegi üritust, mis on ka ilmselt usutav tulemus.



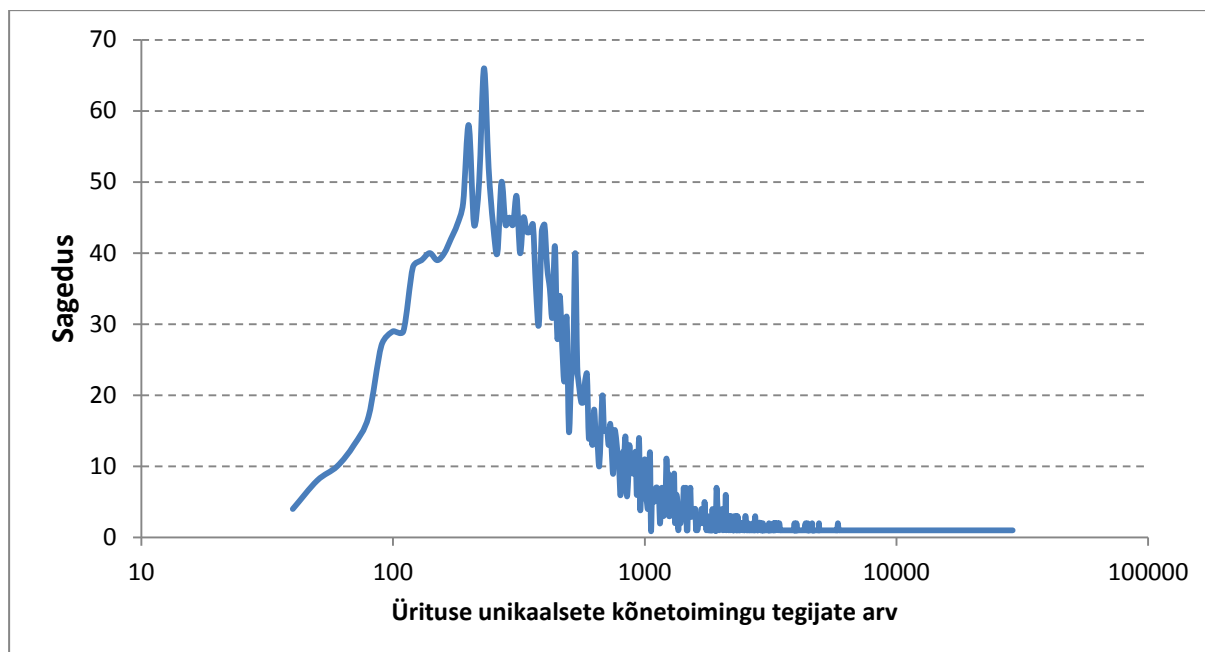
Joonis 11. Tuvastatud ürituste jaotumine ööpäeva lõikes.

Ürituste suurust on võimalik vaadelda mitme näitaja abil. Joonisel 12 on näidatud ürituste kestvuse sageduse jaotust. Ürituse kestvuse aritmeetiline keskmine on 1,9 tundi. Üle poolte ürituste kestvus on 1 tund.



Joonis 12. Tuvastatud ürituste kestvuste sagedusjaotus.

Enim, 52 tuvastatud üritusel, on 240 unikaalsete kõnetoiminute tegijat (joonis 13). Tuvastatud ürituste unikaalsete kõnetoiminute tegijate mediaan on 433. Tuvastatud üritustel osalenud unikaalsete kasutajate arv erineb kümnetes kordades. Suurim tuvastatud üritus on Robbie Williamsi kontsert 28990 unikaalse kasutajaga.



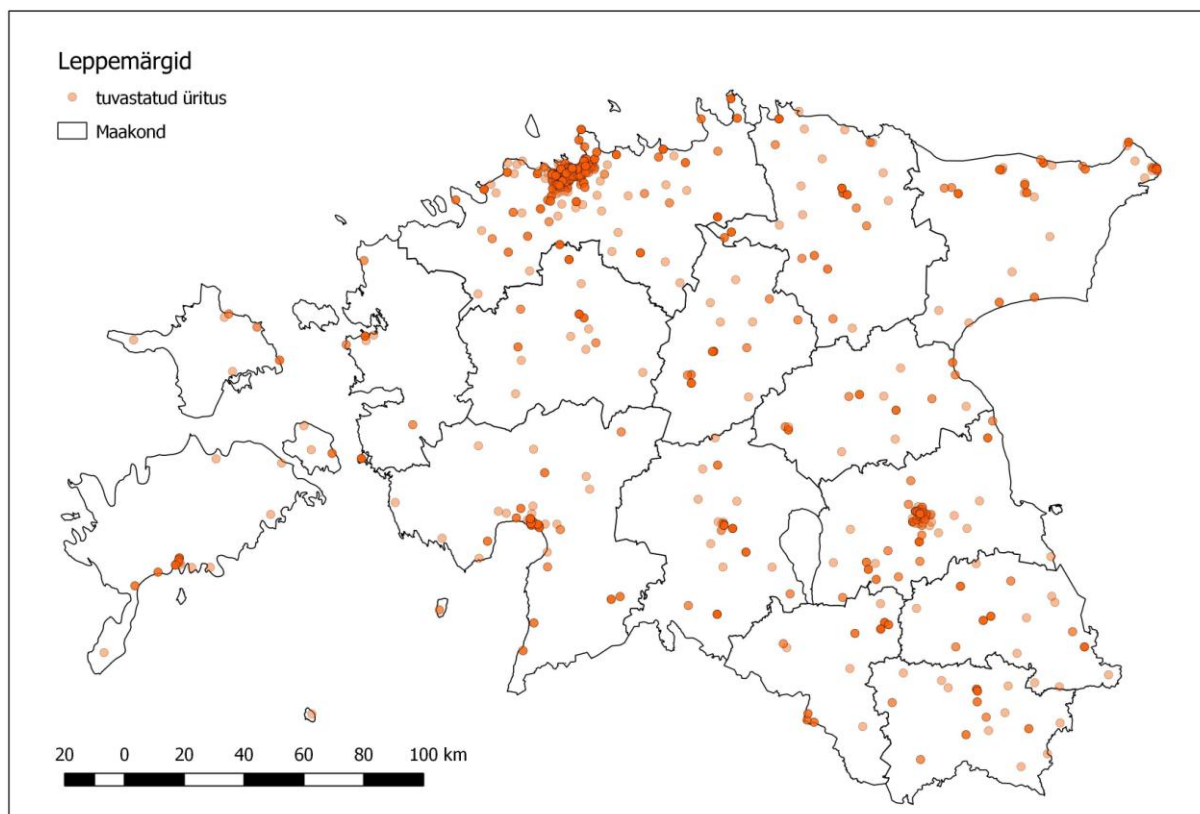
Joonis 13. Tuvastatud ürituste unikaalsete kõnetoimingu tegijate arvu sagedusjaotus.

Tuvastatud ürituste toimumisala pindala varieerus 0,2ha Tallinna kesklinnas kuni Saaremaa rallini, mis kattis 1167 km², mis on umbes pool Saaremaa pindalast. Kõige enam leidis üritusi, mille pindala oli 1 ha ja 10 ha vahel (tabel 1)

Tabel 1. Tuvastatud ürituste toimumisala pindala.

Toimumisala pindala (m ²)	sagedus	osakaal
10 ³ - 10 ⁴	441	14,54%
10 ⁴ - 10 ⁵	983	32,42%
10 ⁵ - 10 ⁶	782	25,79%
10 ⁶ - 10 ⁷	410	13,52%
10 ⁷ - 10 ⁸	361	11,91%
10 ⁸ - 10 ⁹	54	1,78%
10 ⁹ - 10 ¹⁰	1	0,03%
Kokku	3032	100%

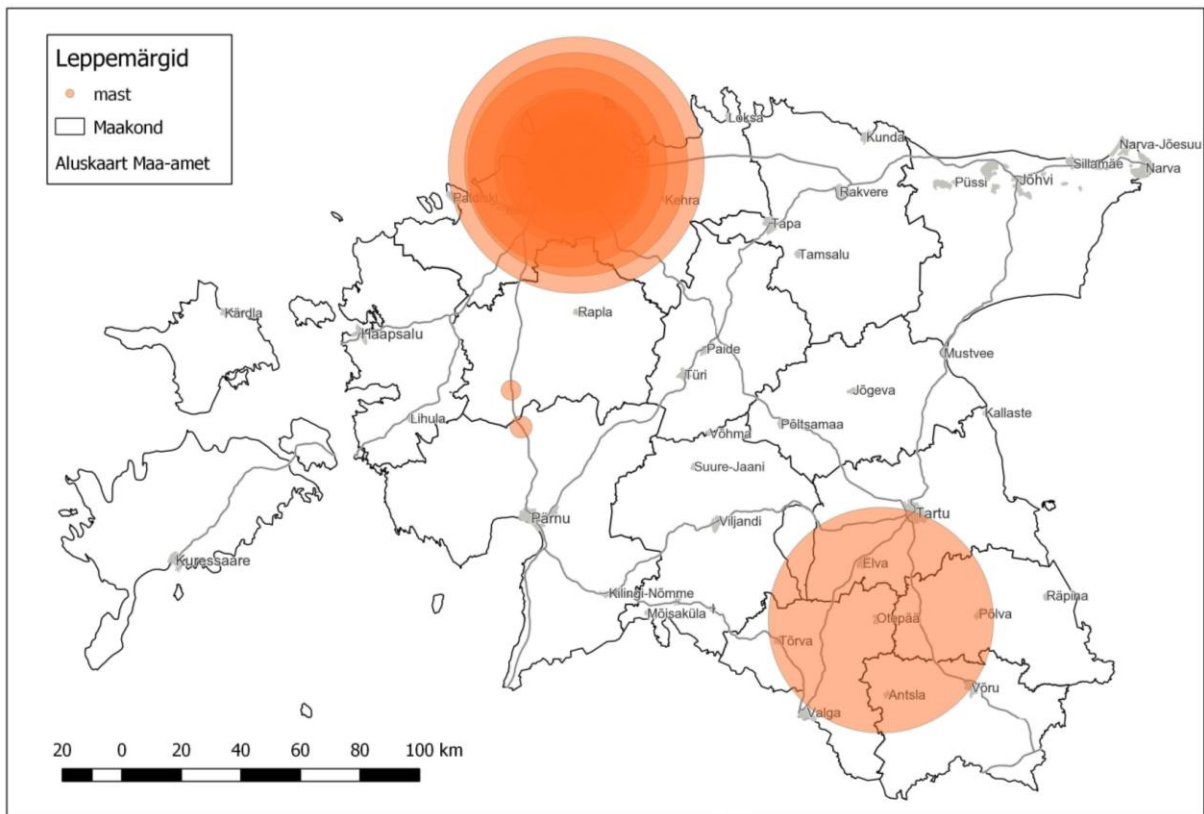
Tuvastatud üritused jagunesid kõikide maakondade vahel. Eristada saab Eesti suuremaid linnasid, kus vastavalt suuremale elanike arvule toimus ka rohkem ja suuremaid üritusi. Kõige rohkem üritusi toimus Tallinnas ja selle lähiümbruses (joonis 14)



Joonis 14. Tuvastatud ürituste paiknemine Eestis.

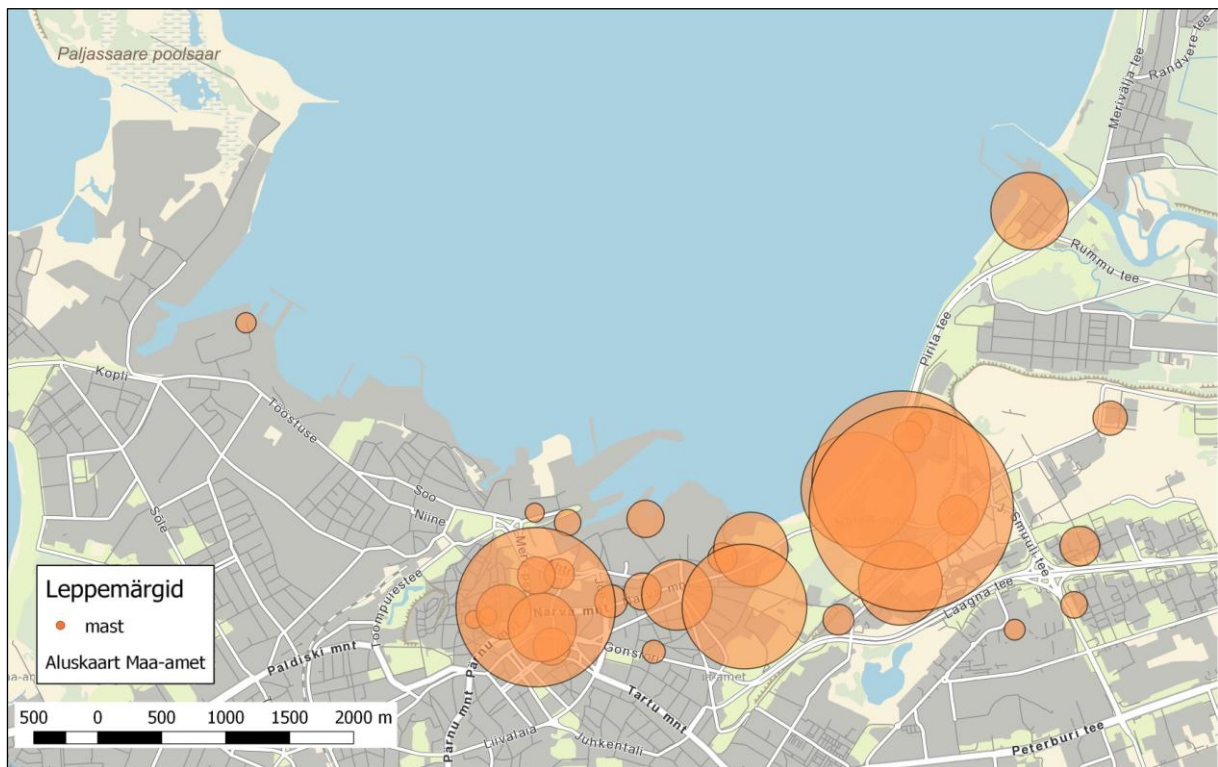
3.2 Aja ja asukoha tuvastamine

Robbie Williamsi kontsert toimus 20. augustil Tallinna lauluväljakul. Korraldajate plaanide järgi avatati väravad kell 16 ja soojendusesineja alustas kell 19:30. Ürituse orienteeruv lõpp on kell 23 (First Baltic International 2013). 19. augustiks oli üritusele müüdud üle 55000 pileti (Toomemets 2013). Lisaks Robbie Williamsi kontserdile toimus Kadrioru pargis kella 15 ajal laulupäev, mis kestis 3 tundi. Üritusel osales ligikaudu 1000 koori ja oskestri liiget (Tallinn 2013). Lisaks toimus Kadriorus, presidendi kantselei roosiaias, Eesti riigi iseseisvuse taastamise 22. aastapäeva vastuvõtt (Vabariigi Presidendi Kantselei 2013). Enamus ürituse osana tuvastatud mastidest asusid Laululava ümbruses ja Tallinna kesklinnas, kuid sinna on kaasanud ka kaks masti Tallinna–Pärnu–Ikla maanteel (joonis 15), kus oli tavapärasest enam välismaalaste kõnesid kella 15 ajal, kes hiljem külastasid Robbie Williamsi kontserdit. Lisaks on üks masti asukoht paigutatunud Otepäele. Tegemist ei ole grupeerimise veaga vaid antud masti saab transportida ühest kohast teise ja käesolev uurimustöö ei osanud selliste nõ liikuvate mastidega algselt arvestada.



Joonis 15. Robbie Williamsi kontserdi ürituse mastide asukohad ja unikaalsete kõnetoimingu tegijate arv tuvastatud ürituse tundidel masti kohta.

Tallinnas on näha kolm suurimat kasutajate arvuga mastide gruppi: Lauluväljak, Kadriorg ja Viru väljak (joonis 16). Lauluväljakul toimus kontsert, mis on ka kõige suurema kasutajate arvuga mast ürituste tundide raames. Pärastlõunal toimus Kadriorus laulupäev. Selle lõppedes suundusid paljud Lauluväljakule kontserdile. Kesklinnas on tavapärasest enam kasutajaid, sest paljud välismaalased tulid läbi Viru väljaku kontserdile ning peale kontserdi liiguti edasi Vanalinna. Peale Lauluväljaku on üritusse kaasatud mastid Lasnamäel, need on põhjustatud liiklusummikutest, mis tekkisid peale kontserdi lõppemist kontserdil osalenutest või kasutati neid võrgu ülekoormuse tõttu. Öötundidel on vähem helistajaid (joonis 4) ning tavapärasest erinevad liikumised tulevad paremini esile. Nii Kopli poolsaarel kui Pirita sadamas on tuvastatud ürituse osad. See tuleneb sellest, et nii laevaga tulles kui ka minnes on need mastid, mille külge Tallinna lahes olles ühendutakse. Kuna paljud lahkusid samal õhtul laevaga, siis on ürituse juurde kaasatud ka sadamas tavapärasest rohkem välismaalaste kõnetoiminguid.

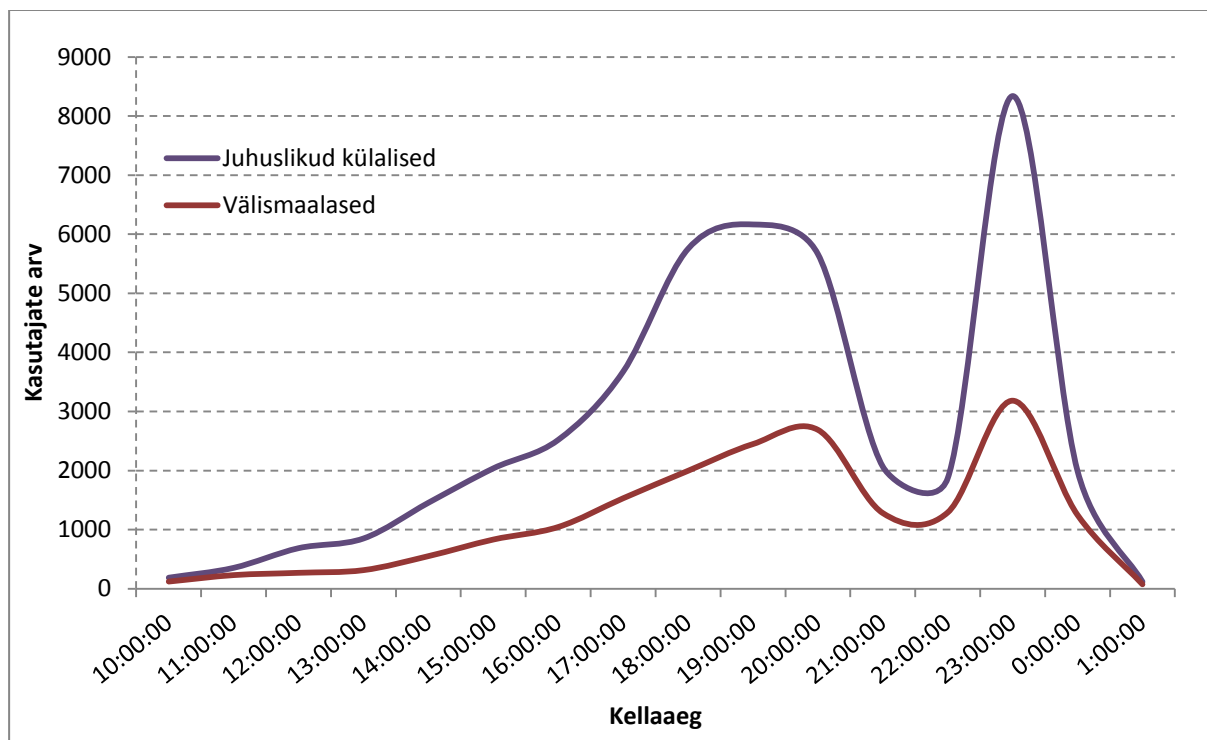


Joonis 16. Robbie Williamsi kontserdi ürituse mastide asukohad Tallinnas ja unikaalsete kõnetoimingu tegijate arv tuvastatud ürituse tundidel masti kohta.

Vaadates ürituse ajal kõne toimingute tegijate ajalise jaotust, on näha, et kõnetoimingute tegijate arv kasvab järsult alates kella 16, millal avati väravad kuni kella 20-ni, millal algas kontsert (joonis 17). Kontserdi ajal vähenes kõnede arv järsult, üle poole võrreldes kontserdi

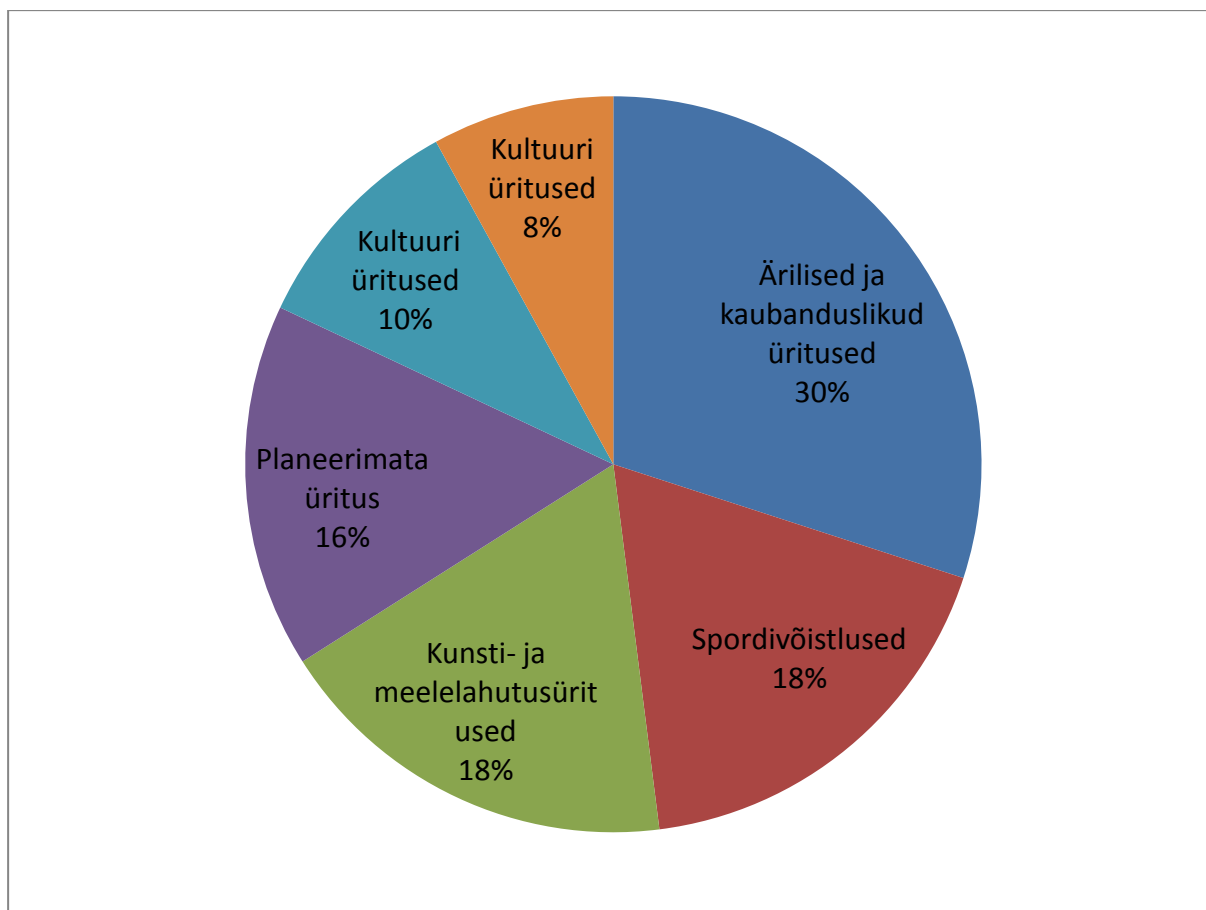
algusega. Ürituse jooksul tehti kõnesid kõige enam peale ürituse lõppu kell 23, kui inimesed hakkasid uuesti liikuma.

Üritusel osalejate arvu ennustamiseks kasutatakse korrelatsioone Kõnetoimingu tegijate arvu ja Antud piirkonnas viibivate inimeste vahel. Tavapärane suhe kõnetoimingu tegijate ja tegelikult antud kohas ja ajal viibivate inimeste vahel on 3,2 inimest ühe kõnetoimingu tegija kohta. Antud näitaja tuli Robbie Williamsi kontserdi üritusel 1,9 inimest ühe kõnetoimingu tegija kohta. Erinedes seega tavalisest kõnede arvust inimese kohta oluliselt.



Joonis 17. Robbie Williamsi kontserdi ennustuse ja tegelike unikaalsete kasutajate jaotumine tundide kaupa.

Kõige enam tuvastati ärilisi ja kaubanduslikke üritusi, mille alla kuuluvad messid, laadad ja ostukampaaniad, nagu osturalli Tallinna kaubamajas ja hullud päevad Stockmannis. Sellele järgnesid spordiüritused ning kunsti- ja meelelahutusüritused, näiteks kontserdid. 16% juhtudest ei tuvastatud kindlat üritust ja liigitati planeerimata ürituseks. Selle alla kuulusid liiklusummikud, kruiisilaevade saabumised, poodlemised enne pühi ja kõige enam tuvastatud üleriigilised üritused, näiteks aastavahetus ja volbriöö. 50 suurima unikaalsete osalejatega ürituse seas ei leidunud ühtki haridus ja teaduslikke, poliitilisi ja riiklikke, rekreatsioonilisi üritusi ega ka erapidusid.



Joonis 18. Tuvastatud ürituste jaotumine liigiti.

4. Arutelu

4.1 Ürituste tuvastamise täpsus

Hetkel ei leidu täielikku ürituste andmebaasi, kus oleks kirjeldatud ürituse kestvus, toimumisala ja osalejate hulk, millega oleks võimalik võrrelda antud meetodil uuritud üritusi. Töös käsitletava meetodiga ei ole võimalik tuvastada, kas tavalisest erinev kõne aktiivsus on seotud planeeritud või planeerimata üritusega. Kuid võimalik on tuvastada mõlemaid ürituse liike. Planeeritud üritustel ei kattu tuvastatud ürituse algusaeg tegeliku ürituse algusajaga. Üheks peamiseks põhjuseks on see, et inimeste tavapärasest erinev käitumine ja liikumine algab juba üritusele minemisel ja jätkub ürituselt tulemisel. See muudab ürituste kestvuse pikemaks ning ürituse toimumisala suuremaks kui ürituse planeerimisel arvestatakse.

Kaasates ürituse juurde ka sinna minemise ja tulemise mast-tunnid, milles on oluliselt suurem osakaal üritusel mitte osalenuid kui ainult ürituse planeeritud toimumisajal, vähendab kõnetoimingute arvu ja sellest lähtuvate analüüside täpsust. Seda oleks võimalik teha algoritmi täiendamisega.

Ürituste tuvastamise täpsust muudab kehvemaks ka mastide erinev tihedus piirkonniti. Näiteks Saaremaa ralli üritusel on maste, millel leviala maismaal on veerand kogu eeldatavast levialast ning ülejäänud katavad Läänemerd, kus üritus ei toimu, kuid lisab ürituse geograafilisele ulatusele suure osa juurde. Kesklinnades, kus mastide võrk on tihe, mastide levikualad kattuvad väga suurel määral. Kuigi telefon lülitab ennast tavaliselt suhtlema lähima mastiga, siis üritustel, kus võrk on ülekoormatud, võib telefon ühenduda ka mõne kaugema mastiga. Kokkuvõtteks on ürituste toimumisala suurust raske hinnata ja võrrelda.

4.2 Ürituste tuvastamisel tekkivad vead

Mobiilpositsioneerimise andmetelt on võimalik tuvastada vaid neid üritusi, kus on tavapärasest rohkem inimesi ja seega kasvab ka kõnetoimingute ja kasutajate arv piirkonnas või muutuvad ürituse tõttu inimeste mobiili kasutusharjumused. Näiteks inimesed helistavad kohtadest, kust nad igapäevaselt ei helista või teevad kõnetoiminguid tihedamini. Üritustel, kus telefoni kasutus langeb (näiteks mõned kontserdid ja etendused), sest palutakse välja

lülita mobiiltelefonid, ei kasva ka kordades suurema rahvaarvu korral unikaalsete kõnetoimingu tegijate arv piisaval määral, et lugeda seda tavapärasest erinevaks.

Vead tekivad veel klasterdamise metoodikast lähtuvalt. Peamiseks klasterdamise alusel tehtavaks ürituse tuvastuse veaks on väga tihedalt omavahel seotud üritused, mis klasterdamisel moodustavad ühise ürituse. Samas nõrgalt seotud üritusest võib tekkida mitu tükki, kuna need pole omavahel piisavalt seotud. Näiteks enne Robbie Williams kontserdi toimus Kadriorus laulupäev, millest paljud osavõtjad liikusid edasi lauluväljakule. Need kaks üritust olid omavahel väga tihedalt seotud ning grupeeriti üheks ürituseks. Samuti Saaremaa Ralli kaks päeva seoti üheks ürituseks kokku samas kui kolme päevane Viljandi Folk grupeerus iga päev eraldi üritusteks. Ürituse erinevad osad, eriti mitmepäevased üritused, võivad omavahel olla väga erinevate ühiste kasutajate arvuga seotud. Samas on DBSCAN hea müra eemaldamises. Võrreldes K-mean klasterdamis algoritmiga ei grupeerita kõiki üritusi gruppidesse, jättes välja Peale selle võimaldab DBSCAN grupeerida väga erineva kujuga üritusi.

Ürituse tuvastamisel tekivad vead kui toimuvad mastivõrgustikus muutused, mastide asukohti on vahetatud, juurde lisatud või ära võetud. Selle alla käivad ka mastide võrgust väljalangemised elektrikatkestuste ajal või hooldustöödeks. Muutused ei kajastu vaid ära langeva või juurde pandava masti andmerekas, vaid ka ümberkaudsetes mastides, kuhu suunatakse kõned ümber. Sellised muudatused võrgustikus suurendava tavapärasest erinevaid kasutajate arvu. Andmevigade vältimiseks võeti kasutusele juhuslike kõnede ja välimaalaste kõnede ennustamisel ± 5 nädala samal nädalapäeval ja tunnil antud masti mediaan. Mediaan aitab vältida ühekordseid tõuse ja langusi andmegraafikus.

Suure rahvaarvuga üritused tulevad grupeerimise juures hästi välja, sest nad eristuvad selgelt tavalisest rütmist. Väiksemate ürituste puhul võivad jääda osad ürituse osad potentsiaalsete ürituste juurde arvamata ning ürituse osade omavahelised sarnasused võivad olla väga nõrgad.

Antud metoodika parandamiseks tuleks uurida, kas masinõppe abil oleks võimalik tuvastada ürituse liike ja millise täpsusega. Lisaks aitaks täpsema ja paremate tulemuste saamiseks kõigepealt andmevigadest puhastamine. Antud metoodika ei keskendu andmevigade otsimisele, samas paljud andmevead on antud metoodikaga võimalik tuvastada. Antud töö raames ei uuritud üritustel osalenud inimeste arvu ja kõnetoimingute omavahelist korrelatsiooni eri rahvuste ja piirkondade vahel. Robbie Williamsi kontsert näitas, et see on tavapärasest tunduvalt erinev ja vajab täpsemat uurimist. Korrelatsiooni leidmine erinevate piirkondadel ja rahvustel aitaks täpsustada üritusel osalenud rahvaarvu ja rahvuslikku

jaotumist. Korrelatsioon osalejate arvu ja kõnetoimingute vahel on sõltuv ürituste liigist. Võimalik oleks ka täpsustada ürituse geograafilist paiknemist, leides alad, kus inimesed kõige tõenäolisemalt oma aega veedavad.

Töös toodud metoodika algoritm on optimeerimata ning seda ei ole võimalik kordades suuremate andmemahtudega mõistliku ajaga teostada. Selleks tuleks põhjalikumalt uurida, kuidas oleks kõige suuremat ressursi nõudvaid protsesse võimalik teostada vähemate tehete abil.

5. Kokkuvõte

Käesoleva uurimistöö eesmärgiks oli koostada algoritm, millega on võimalik tuvastada üritusi passiivsete mobiilpositsioneerimise andmete abil ning kirjeldada algoritmi töö tulemusi.

Algoritmi loomiseks kasutati Eesti 2013. aasta mobiilpositsioneerimise andmeid, mis oli kogutud OÜ Positium LBS poolt. Algoritmi loomisel kasutati ankurpunktide arvutamise metoodikat (Ahas et al 2010), kus Eesti mobiilsideoperaatori kasutajate tegevusruum kirjeldati ning leiti need kohad, mis pole igapäevase tegevusruumi osad. Algoritmi loomisel leiti potentsiaalsete ürituste mastide ja tundide klastrid, kus kasutajate arv erines olulisel määral tavapärasest.

Kokku tuvastati 2013. aastal 3032 üritust. Tuvastatud ürituste seas oli erineva kõnetoimingu tegijate arvu, pikkuse, algusaja ja geograafilise suurusega ürituse. Üritused paiknesid üle Eesti jälgides rahvastiku paiknemist. Ürituste toimumisala pindala varieerus suuresti alates 0,001 km² kuini 1000 km². 50 suurimale unikaalse kõnetoimingu tegijate arvuga üritust prooviti seostada üritustega. Vaid 8 üritusele ei leitud vastet. Kõige enam tuvastati ärilisi ja kaubanduslikke üritusi (30%). Sellele järgnesid võrdselt spordiüritusi ja kunsti- ning meelelahutusüritused. 50 suurima unikaalsete osalejatega ürituse seas ei leidunud ühtki haridus- ja teaduslikke, poliitilisi ja riiklikke, rekreatsioonilisi üritusi ega ka erapidusid. Selle põhjuseks võib-olla, et antud üritused ei olnud 2013. aastal Eesti suurimate ürituste seas ja antud liiki üritustel tavapärasest väiksema telefonikasutuse.

Edaspidistes uurimustes võiks tulemusi võrrelda ürituse osalejate arvu mõne teise meetodiga, et saaks luua korrelatsioonid unikaalsete kõnetoimingute tegijate ja ürituste vahel. Korrelatsioonide abil oleks võimalik hinnata ürituse osalenute arvu ja välismaalaste osakaalu üritusel, mis antud töös jäi kajastamata.

Töö tulemusena võib kinnitada, et antud algoritmiga on võimalik tuvastada erineva pikkuse, geograafilise suuruse ja liiki üritusi. Leida ürituse algus- ja lõpuaeg ning mõõta ürituste geograafilist ulatust.

Tänuavaldused

Soovin tänada juhendajaid Erki Saluveeri mitmekülgse abi ja nõuannete ning pühendatud aja eest. Tänan ka Anto Aasat ja Positium LBS töötajaid mitmete tähelepanekute ja täienduste eest töö koostamisel ja vormistamisel. Veel tahan tänada igati toeks olnud sõpru ja peret.

Taavi Pipar

Summary

Events detection from mobile positioning data

Taavi Pipar

The purpose of this research was to create an algorithm which is able to detect events from passive mobile positioning data and describe the results. The algorithm should find events start and end time and event geographical size. The produced algorithm was based on Estonian 2013th mobile positioning data collected by Positium LBS. University of Tartu Anchor points calculation (Ahas et al 2010) was used to divide Estonian mobile operator calls into two: an anchor point and occasional visits. Calls were grouped by sites and hours. From hour and site clusters were selected potential events where the number of Visitors and occasional Estonian visitors was significantly bigger than normally. Potential events were clustered using DBSCAN where the similarity was defined by shared users.

There were 3032 events detected in 2013 with this algorithm. Detected events had wide variation of length, start time, number of callers and geographic size. Events were located all over Estonia following the density of population. The size of event location varied widely from 0,001 up to 1000 square kilometres. For top 50 events with highest number of callers was tried to find the events type. Only for 8 events was not find type mainly because they was unplanned events. The most frequent type was business and trade event (30%). This was followed by equally sport competition and arts and entertainment events. Among the top 50 events there was not found political and state, educational and scientific and private events. The reason may be that these types of events are not that big to be in top 50.

In the future studies the results could be compared with another method to find different correlations between event types and country of origin. As a result it is confirmed that the algorithm is able to detect different size and types of events. The algorithm also finds events start and end time and event geographical size.

Kasutatud kirjandus

- Ahas, R., Silm, S., Järv, O., Saluveer, E., Tiru, M.** (2010). Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones. *Journal of Urban Technology*, 17(1), 3-27.
- Ahas, R., Aasa, A., Mark, Ü., Pae, T., Kull, A.** (2007). Seasonal tourism spaces in Estonia: Case study with mobile positioning data. *Tourism Management*, 28(3), 898-910.
- Bonné, B., Barzan, A., Quax, P., Lamotte, W.,** (2013). WiFiPi: involuntary tracking of visitors at mass events. IEEE 14th international symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), Madrid, Hispaania.
- Getz, D.,** (2008). Event tourism: Definition, evolution, and research. *Tourism Management*, 29, 403–428.
- Getz, D.,** (2005). Event Management and Event Tourism. New York-Sydney-Tokyo: Cognizant Communication Corporation 439pp.
- Gursoya, D., Kimb, K., Uysalc, M.,** (2004). Perceived impacts of festivals and special events by organizers: an extension and validation. *Tourism Management*, 25, 171-181.
- Jago, L., Shaw, R.,** 1998. Special Events: A Conceptual and Differential Framework. *Festival Management & Event Tourism*, 5(1/2), 21-32.
- Mossberg, L. L.,** (2000). Evaluation of Events: Scandinavian Experiences. New York: Cognizant Communication Corporation. 161pp.
- Nilbe, K.,** (2011). Turismiürituste tagamaad ja nende ulatust mõjutavad tegurid Eesti turismiürituste näitel.
- Nilbe, K.,** (2008). Turismiürituste geograafilise tagamaa hindamine mobiilpositsioneerimise andmetega Põlva ökofestivali ja Hauka laada näitel.
- Shoval, N., McKercher, B., Ng, E., Birenboim, A.,** (2011). Hotel location and tourist activity in cities. *Annals of Tourism Research*, 38(4), 1594-1612.
- Versichele, M., de Groote, L., Bouuaert, M.C., Neutens, T., Moerman, I., Van de Weghe, N.,** (2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management*, 44, 67-81.
- Öztaysi, B., Baysan, S., Akpınar, F.,** (2009). Radio frequency identification (RFID) in hospitality. *Technovation*, 29(9), 618-624.

Internetiallikad:

Tallinn (2013). Laulupäev <http://www.tallinn.ee/laulupaev> (15.05.2014).

Vabariigi Presidendi Kantselei (2013) President Ilves iseseisvuse taastamise aastapäeval: Eesti edu on alati seisnenud uuesti alustamises <http://president.ee/et/meediakajastus/pressiteated/9322-2013-08-20-13-56-01/index.html> (15.5.2013).

First Baltic International. (2013) Robbie Williamsi Tallinna kontserdi külastaja meelespea <http://www.fbi.ee/robbie-williamsi-tallinna-kontserdi-kulastaja-meelespea/> (15.5.2014).

Toomemets, K., (2013). Robbie kontserdi korraldaja palve: võtke kaasa lipud ja jätkke kodus tuled põlema <http://www.ohtuleht.ee/537652/robbie-kontserdi-korraldaja-palve-votke-kaasa-lipud-ja-jatke-kodus-tuled-polema> (15.5.2014).

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Taavi Pipar (sünnikuupäev: 27.06.1992),

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
„Ürituste automaatne tuvastamine passiivse mobiilpositsioneerimise andmetelt“,

mille juhendaja on Erki Saluveer,

1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, **19.05.2014**